Purdue Univ.

NsG-28-59

56p.

UNPUBLISHED PRELIMINARY DATA

N63 23815

CODE-1

NASA CR 52226

OTS PRICE

XEROX $ [redacted]

MICROFILM $ [redacted]

OPTIMUM NONLINEAR CONTROL OF A SYSTEM

WITH TWO INPUTS AND ONE OUTPUT

A Thesis

Submitted to the Faculty

of

7243009 Purdue ~~University~~ Lafayette, Ind.

by

Kiyohisa Okamura (Ph.D. Thesis)

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

June 1963 56 p refs

(NASA Grant NsG-28-59)

(NASA CR-52226) OTS: $5.60 ph, $1.88 mf

ACKNOLEDGEMENT

I would like to express my deepest appreciation to my major professor, Dr.Rufus Oldenburger, for his encouragement, guidance and advice during my graduate study.

I wish to extend greatful thanks to Professors E.H.Gamble, K.S.Fu and A.Sinclair for their assistance.

The valuable advice and encouragement of Dr.C.Rezek, Dr.R.E.Goodson, Dr.R.Kohr, Mr.F.D'Souza, Mr.R.Boyer and my fellow graduate students are acknoledged with thanks.

I am indebted to National Aeronautical and Space Administration for the financial assistance that made possible my graduate study at Purdue University.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

ABSTRACT

Okamura, Kiyohisa. Ph.D., Purdue University, June 1963.

Optimum Nonlinear Control of a System with Two Inputs and One Output.

Major Professor: Rufus Oldenburger

23815

Much work has been done in the field of optimum nonlinear control of multivariable systems. This effort has been largely mathematical and programming actual optimum control schedules is still difficult and laborious. More direct methods to achieve an optimum controller for multivariable systems as a function of the state variables may be quite useful in many cases. This thesis proposes the extension of Oldenburger's approach to optimum nonlinear control of single variable systems to that of a simple class of multivariable systems. The class next in complication to single input-single output systems may be the one where the system has two inputs and one output. This class may be treated as a single input-single output system except where one of the inputs and the time derivative of the other are bounded. This class of systems is to be optimized such that the time duration of the transient is minimized as well as the maximum over or under swing, etc. After the transient vanishes both inputs are to be brought to their reference values as quickly as possible. A controller is designed which yields an optimum transient. The controller uses a check-decision process which does not arise in a single input-single output system. The approach developed here is direct and straightforward rather than abstractly mathematical.

AUTHOR

INTRODUCTION

This thesis concerns optimum nonlinear control of a simple class of multivariable systems. The control system is composed of a plant to be controlled and a controller which optimizes the response of the controlled quantity of the plant according to specified criteria. This type of control has been treated by many engineers and mathematicians.

In 1944 R.Oldenburger[1] derived a control scheme to obtain the optimum transient of the system with a bounded input. He did this while he was studying an aircraft engine-propeller system. D.McDonald[2] in 1950 published a paper on optimum nonlinear control of second order systems with bounded input. This was followed by papers from 1951 to 1959 by A.Hopkin[3], Uttley and B.H.Hammond[4], I.Flügge-Lotz[5], L.F.Kazda[6], T.M.Stout[7], and others. L.M.Silva[8], T.Bogner[9] both in 1954 and S.S.L.Chang[10] in 1955 treated systems of third and higher orders. Nonlinear systems were covered by R.Oldenburger, J.C.Nicklass and E.H.Gamble[11] in 1961.

In 1953 D.W.Bushaw[12] published a paper concerned with a treatment on determining the switching for optimum control. In 1956 R.Bellman, I.Glicksberg and O.Gross[13] applied Bellman's dynamic programming technique to optimum nonlinear control systems. In 1957 L.S.Pontryagin[14] presented a new principle applicable to multivariable control systems with bounded inputs. More detailed discussion about Pontryagin's maximum principle was given by L.I.Rozonoer[15]. J.P.LaSalle[16] in 1960 published a proof of the existence theorem for optimum transients. In 1962 several papers on optimum nonlinear multivariable control systems were published by

B.Friedland[17], L.Markus and E.B.Lee[18], E.R.Rang[19], Yu-Chi Ho[20], G.Boyadjieff, D.Eggleton, M.Jackues, H.Sutabutra and Y.Takahashi[21] and others. The works cited above are largely mathematical and programming actual optimum control schedules by these methods is difficult and laborious. Furthermore, practical applications have not been extensively treated. This thesis proposes the extension of Oldenburger's approach[1] to optimum nonlinear control of single variable systems to that of a simple class of multivariable control systems where a practical engineering example is treated. The approach developed here is direct and straightforward rather than abstractly mathematical.

Oldenburger treated systems with one controlling variable and one controlled quantity where the controlling variable or its time derivative is bounded. The systems next in complication to single input-single output systems are those which have one controlled quantity and more than one controlling variable. However, for the case of a system where each of the controlling variables but none of the time derivatives are bounded or, conversely, where the time derivative of each controlling variable but none of the controlling variables are bounded, the treatment is the same as for single input-single output systems. The reason for this is given as follows. If each controlling variable is limited, the sum of the controlling variables can be made to take on any value between two limits where one of these limits is the sum of the upper bounds and the other the sum of the lower bounds of the controlling variables. Thus the sum of the controlling variables is treated as a single input. The same argument holds for the case when the time derivative of each controlling variable is bounded but not the controlling variables themselves.

Multiple input-single output systems which are not equivalent to single input-single output systems as explained above are those which have inputs composed of a mixture of bounded variables and variables with bounded derivatives. Dividing the inputs into two sums, one of which contains the variables whose derivatives are bounded and the other containing the remaining variables which are limited. The above systems may be treated as dual input-single output systems. This is the case with which this thesis is concerned.

The control system treated in this thesis has a physical interpretation as follows. The controlled quantity is the deviation in the level of the surface of a liquid in a tank from a reference value. One controlling variable is the rate at which liquid flows out of the tank through a pump. The other is the position of a valve controlling flow to a tank as shown in Figure 1. The acceleration and deceleration of the pump connected to the motor are limited whence the rate of change of flow out of the tank is bounded. The position of the valve is usually limited. We have a case where one of the controlling variables and the time derivative of the other controlling variables are bounded. Our object is to bring the controlled quantity to zero in an optimum sense (the time duration of the transient is minimized as well as the maximum over or under swing, etc.) after the disturbance dies out. One of the controlling variables can be made to be "bang-bang", but the other cannot.

The time optimal control of this system is not always uniquely determined. We consider three hypotheses to find the optimum control schedule for which we derive two kind of control functions. Since the controlled plant has two inputs and one output, the output can be set equal to zero while the inputs are not zero. Hence we consider the control schedule, after the controlled quantity becomes zero, in which both controlling variables are brought to zero in minimum time while the controlled quantity remains zero. Thus the control of the system consists of two stages. In the first stage the controlled quantity is to be brought to zero. In the second stage the controlling variables are brought to zero. The programs to achieve the optimum control schedules for the above two stages are designed from practical considerations. The controller must go through a check-decision process in addition to examining signs of switching functions. This does not arise in the case of single input-single output systems. It is mathematically proved that all transients based on the above control schedules are optimum.

The closed loop system with the optimum controller was tested on an electronic analog computer. Relays were used in logical circuits which realize the optimum control schedules. Good agreement was found between theory and analog simulation results.

LIQUID SURFACE CONTROL SYSTEM

The following differential equation is associated with the tank in Figure 1

$$A \frac{d\,h(t)}{d\,t} = q_i(t) - q_o(t) \tag{1}$$

where

$h(t)$ = level of liquid surface (deviation from reference value)

$q_i(t)$ = rate of incoming flow (deviation from reference value)

$q_o(t)$ = rate of outgoing flow (deviation from reference value)

A = constant area of liquid surface

The rate of incoming flow is taken proportional to the position of the valve which is limited, i.e.

$$q_i(t) = \alpha \ell(t) \tag{2}$$

$$|\ell(t)| \leq L \tag{3}$$

where

$\ell(t)$ = position of the valve (deviation from reference value)

α, L = positive constant

The rate $q_o(t)$ of outgoing flow is proportional to the revolution per unit time, i.e. rpm of the motor. Hence it follows that

$$q_o(t) = -\beta n(t) \tag{4}$$

where

$n(t)$ = revolution per unit time of the pump and motor (deviation from reference value)

β = positive constant

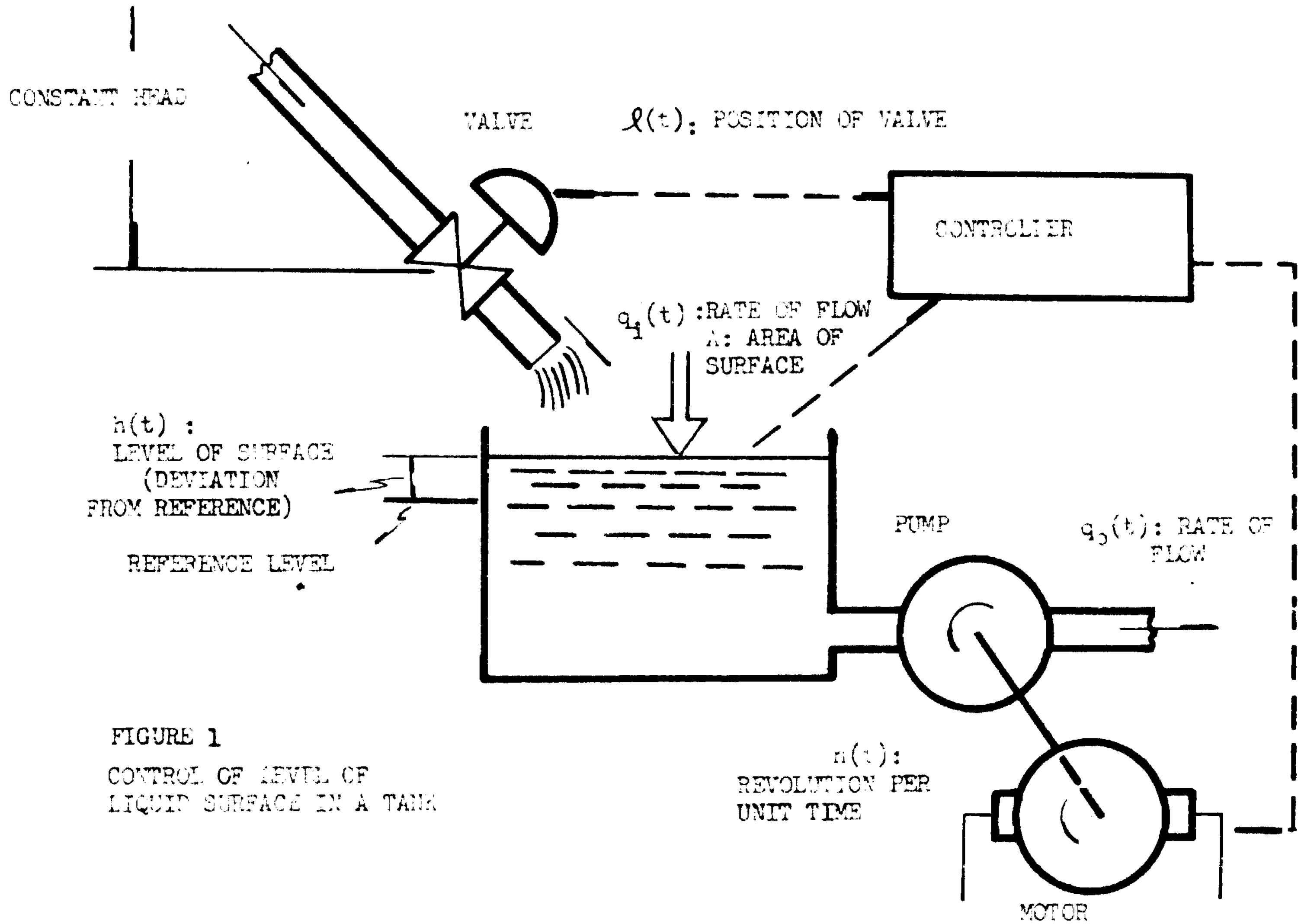


FIGURE 1

CONTROL OF LEVEL OF
LIQUID SURFACE IN A TANK

The angular acceleration and deceleration of a motor connected to a pump are usually limited. It is assumed that

$$\left| \frac{d\,n(t)}{d\,t} \right| \leq N \tag{5}$$

where N is a positive constant. It is also assumed that $q_i(t)$ and $dq_o(t)/dt$ can arbitrarily be made to take on any instantaneous values in the range $-\alpha L \leftrightarrow \alpha L$ and $-\beta N \leftrightarrow \beta N$ respectively. If $h(t) = dh(t)/dt = 0$, the system is said to be in the equilibrium state. This does not imply, however, that the controlling variables are at their reference values, namely $n(t) = \ell(t) = 0$. In this thesis, the system is said to be in the reference state if $h(t) = n(t) = \ell(t) = 0$.

NORMALIZED SYSTEM

We normalize Relations (1), (3) and (5) as

$$Y'(T) = X_1(T) + X_2(T) \tag{6}$$

$$|X_1'(T)| \leq 1 \tag{7}$$

$$|X_2(T)| \leq K \tag{8}$$

by introducing the following dimensionless quantities

$$Y(T) = \frac{A\,h(t)}{N\,\bar{t}^{\,2}}, \quad X_1(T) = \frac{n(t)}{N\,\bar{t}}, \quad X_2(T) = \frac{\alpha\,\ell(t)}{N\,\bar{t}}$$

$$K = \frac{\alpha L}{\beta N\,\bar{t}}, \qquad \frac{t - t_0}{\bar{t}} \tag{9}$$

where

t_o = arbitrary time origin

$\bar{t}$ = arbitrary time scaling factor (may be unit time)

and a prime stands for the derivative with respect to time T.

STAGES OF OPTIMIZATION

We consider two stages which the system will take. In Stage 1 the controlled quantity $Y(T)$ is not identically zero. In Stage 2 the controlled quantity $Y(T)$ is identically zero but not both $X_1(T)$ and $X_2(T)$ are zero. In Stage 1 a control is said to be optimum if the time duration and maximum over (or under) swing of the transient are minimized. If the above control is not unique, the control is said to be optimum when the area between the transient trajectory and the T axis is made as small as possible while the time duration and over (or under) swing of the transient are kept at their minimum values. In Stage 2 a control is said to be optimum if both $X_1(T)$ and $X_2(T)$ are made to be zero in minimum time. Our object is to obtain the optimum control both in Stages 1 and 2.

Suppose that

$$-K \leq X_1 \leq K \tag{10}$$

at some time $T = T_R$. It is possible to keep $X_1(T)$ in the Range (10) after $T = T_R$. For example, if we set $X_1' = 0$ from $T = T_R$, we obtain $X_1(T) = X_1(T_R)$ for $T \geq T_R$, where $X_1(T_R)$ satisfies Relation (10). On the other hand, the variable $X_2(T)$ can be made to take on an arbitrary instantaneous value between $-K$ and K. Therefore, if Relation (10) holds, it is possible to obtain the relation $X_2 = -X_1$, or $Y' = 0$. Hence, if the relation

$$Y(T) = 0 \tag{11}$$

is also obtained at the instant $T = T_R$, it is possible to make $Y(T)$ identically zero from the time $T = T_R$. Conversely, if Relation (10) is not satisfied at any instant, we cannot set $X_2 = -X_1$, i.e. $Y' \neq 0$. Hence, the controlled quantity $Y(T)$ cannot be made zero. As shown above, the quantity $Y(T)$ can be made to be identically zero if and only if Relations (10) and (11) are simultaneously satisfied at some time.

We introduce the following functions for Stage 1:

$$\phi(T) = \int_0^T X_1(\lambda)\, d\lambda + Y(0) \tag{12}$$

$$\psi(T) = -\int_0^T X_2(\lambda)\, d\lambda \tag{13}$$

where

$$Y(T) = \phi(T) - \psi(T) \tag{14}$$

It follows that

$$\begin{aligned} \phi'(T) &= X_1(T) && \text{(a)} \\ \psi'(T) &= -X_2(T) && \text{(b)} \\ \psi(0) &= 0 && \text{(c)} \\ |\phi''(T)| &\leq 1 && \text{(d)} \\ |\psi'(T)| &\leq K && \text{(e)} \end{aligned} \tag{15}$$

Relations (10) and (11) are equivalent to the following:

$$-K \leq \phi'(T) \leq K \tag{16}$$

$$\phi(T) = \psi(T) \tag{17}$$

The variable X_2 may change suddenly at any instant and therefore the derivative Y' is not uniquely determined at that instant. Since the value of X_2 may be chosen at will subject to restriction imposed by Relation (8) it is sufficient to know the value of X_1 in order to determine the derivative Y'. Thus the values $\Phi(0)$ and $\Phi'(0)$ determine the initial conditions $Y(0)$ and $Y'(0)$, where the function $\Phi(T)$ and its derivative $\Phi'(T)$ are continuous. Instead of directly treating $Y(T)$ we consider the functions $\Phi(T)$ and $\Psi(T)$ throughout the analysis of transient responses in Stage 1.

OPTIMUM CONTROL IN THE FIRST STAGE

In this section we treat the control of Stage 1. In this stage it is desired to make $\Phi(T) = \Psi(T)$ in minimum time such that the area between the transient response trajectory and the T axis is minimized. Although analog computer results of optimum transients will be shown later in terms of the controlled quantity $Y(T)$, for the sake of clarity, all possible types of transients are explained here in terms of functions $\Phi(T)$ and $\Psi(T)$. There are six major distinct types of transients in Stage 1 which are labeled Transient Types A through F in the discussion below depending on the conditions at $T = 0$. Each major transient type may have two or more sub types of transients. The treatment of the cases where $Y(0) < 0$ is identical to that when $Y(0) > 0$. This is true because the types of optimum transients possible for $Y(0) < 0$ are mirror images about the T axis of those for $Y(0) > 0$. Treatment of the case for $Y(0) = 0$ is given only for $X_1(0) > 0$ since the types of transients for $X_1(0) < 0$ and $Y(0) = 0$ are the mirror images of their counterparts.

In the figures for illustration of Stage 1, the transient time is denoted by T_R, i.e. Stage 1 terminates at the time $T = T_R$. It is proved in the following sections that for an optimum transient in Stage 1 both the controlling variable X_2 and the time derivative of the other controlling variable X_1' must be set at all times at their extremum values. Hence, for such a transient we have

$$\phi'' = X_1' = \pm 1 \tag{18}$$

$$\psi' = \pm K \tag{19}$$

Since the treatment of the optimum transient is identical for $Y(0)>0$ and $Y(0)<0$, we shall develope the optimum control schedule considering that $Y(0)$ satisfies

$$Y(0) \geq 0 \tag{20}$$

The problem is to determine the optimum values for ϕ'' and ψ' at $T = 0$ as functions of the initial conditions. Noting that $\psi(0) = 0$ from Relation (15-c), we find the solution of Equation (19) to be

$$\psi = \pm KT \tag{21}$$

Plotting the above relation in Figure 2, we obtain two trajectories denoted by Δ_1 and Δ_2 corresponding to the positive and negative signs, respectively, which occur on the right hand side of Relation (21). Thus, the trajectories Δ_1 and Δ_2 are the upper and lower boundaries, respectively, of a family of trajectories representing the function $\psi(T)$.

In order to determine the value of $\phi''(T)$ at $T = 0$ we consider the case

$$\phi'' = 1 \tag{22}$$

Plotting Relation (22) for various initial conditions we examine whether or not the above choice is optimum. The examination could be made by choosing $\phi'' = -1$ instead of Relation (22), but as we shall see, Relation (22) represents the proper choice at $T = 0$ for larger class of initial conditions. Referring Relations (12) and (15-a), there results

$$\phi' = T + X_1(0) \tag{23}$$

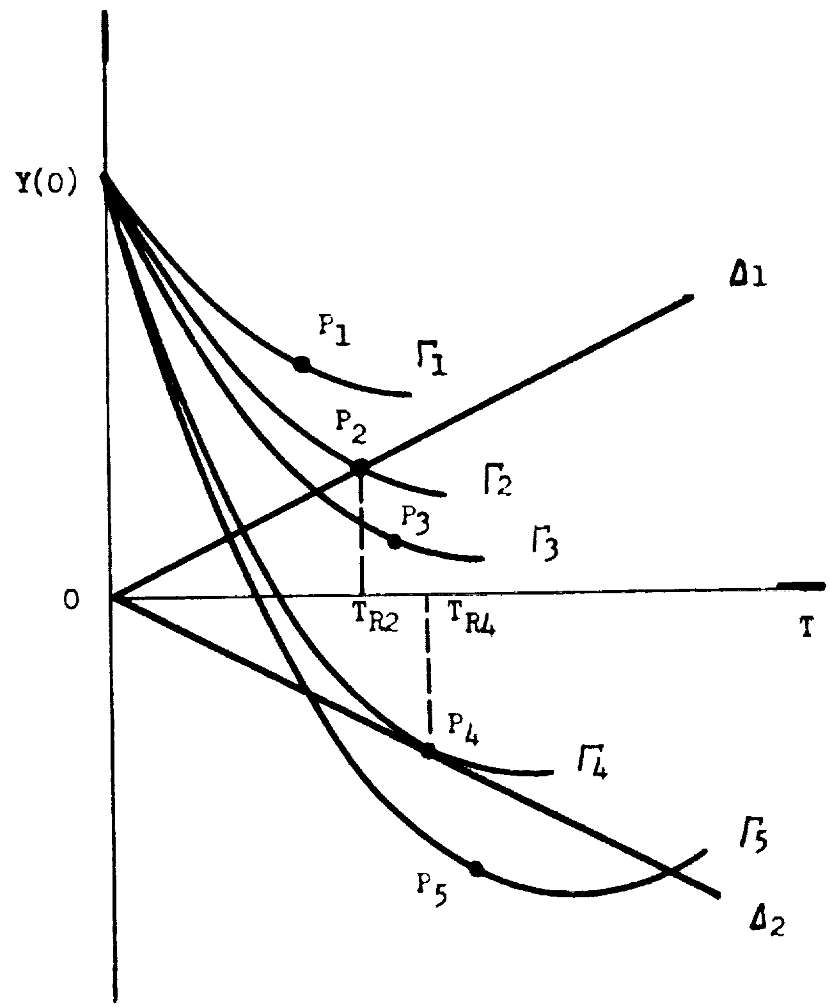


FIGURE 2 A FAMILY OF Φ TRAJECTORIES FOR $\Phi'' = 1$ AND A FAMILY OF Ψ TRAJECTORIES FOR $\Psi' = \pm K$

$$\Phi = \frac{T^2}{2} + X_1(0)\,T + Y(0) \tag{24}$$

Plotting Relation (24) in Figure 2 we obtain a family of trajectories designated by Γ_1 through Γ_5 for $Y(0) > 0$. The case for $Y(0) = 0$ is to be treated later in this section. This family of trajectories is classified as follows. By Relation (23), for each trajectory Γ_i ($i = 1,2,..,5$) there exists a point where the slope of the trajectory is $-K$. This point is denoted by P_i as shown in Figure 2. The point P_i lies above or on the trajectory Δ_1, below or on the trajectory Δ_2, or between the trajectories Δ_1 and Δ_2. There are five cases for $T \geq 0$ which must be considered depending on the location of the points P_i with respect to the trajectories Δ_1 and Δ_2.

Consider a special case represented by the trajectory Γ_2 and its point P_2, where the point P_2 is on the trajectory Δ_1. The functions $\Phi(T)$ and $\Psi(T)$ which correspond to the trajectories Γ_2 and Δ_1 respectively satisfy Relations (16) and (17) at the instant when the point P_2 is obtained. Let the above instant be designated by T_{R2}. Hence, the controlled quantity Y is zero at $T = T_{R2}$ and can be kept identically zero for $T \geq T_{R2}$. By Relation (22) the trajectory Γ_2 is concave up with the slope increasing at the maximum rate. Therefore, any other Φ trajectory with the same initial slope as the trajectory Γ_2 cannot attain the slope equal to or greater than $-K$ before the instant $T = T_{R2}$, i.e. the time duration T_{R2} is the minimum transient time and the trajectory Γ_2 is the unique curve which yields the minimum transient time T_{R2}. Since the trajectory Δ_1 is the upper boundary of a family of Ψ trajectories, there exists no other Ψ trajectory which intersects the trajectory Γ_2 before or at the instant

$T = T_{R2}$. Thus it has been proved that the trajectories Γ_2 and Δ_1 for $T \leq T_{R2}$ uniquely determine the optimum transient in Stage 1 under the given initial conditions. The above type of optimum transient is called Transient Type A.

Another special case concerns the trajectories Γ_4 and Δ_2 where the point P_4 of the trajecotry Γ_4 is on the trajectory Δ_2. Since the trajectory Δ_2 always has slope -K and the slope of Γ_4 at the point P_4 is -K, the trajectory Δ_2 is the tangential line of the trajectory Γ_4 at the point P_4. By the same reason as for Transient Type A, except that the trajectory Δ_2 is here the lower boundary of a family of Ψ trajectories, the trajectories Γ_4 and Δ_2 uniquely determine the optimum transient in Stage 1 for $T \leq T_{R4}$ where the time T_{R4} corresponds to the point P_4. The type of transient obtained above is called Transient Type B.

In the above cases the Φ trajectory corresponding to Γ_2 or Γ_4 is concave up with its slope increasing at the maximum rate and the trajectory corresponding to Δ_1 or Δ_2 is the upper or lower boundary of a family of trajectories. The optimum transient for Stage 1 is obtained without changing the values of X_1' and X_2 for the above two cases. Types A and B are special cases which would not often occur but theoretically important in determining control functions. Types C through F discussed below are the usual types which would generally be found.

If a Φ trajectory with its slope increasing at the maximum rate starts at the same point as the trajectory Γ_2 with greater slope than that of Γ_2, the trajectory Γ_1 in Figure 2 would result. Here the point on Γ_1 denoted by P_1 at which the slope is $-K$ lies above the trajectory Δ_1. Though the case in which the point P_1 exists for $T > 0$ is shown in Figure 2, the point P_1 may be found for $T \leq 0$ in some other cases in which the trajectory Γ_1 has the greater initial slope. It will be proved in the next section that the trajectory Γ_1 is not the optimal one but that the optimal trajectory must be concave down with its slope decreasing at the maximum rate near $T = 0$. To show this let ϕ'' be given by

$$\phi'' = -1 \tag{25}$$

The solution of the above equation is then

$$\phi' = -T + X_1(0) \tag{26}$$

$$\phi = -\frac{T^2}{2} + X_1(0)\,T + Y(0) \tag{27}$$

In plotting Relation (27), two cases must be considered. For the first case, the plot of Relation (27) designated by Γ as shown in Figure 3(a) has slope $-K$ at the point denoted by P which is on or below the trajectory Δ_1 where the trajectory Δ_1 is the same as in Figure 2. Since the trajectory Γ is concave down, it intersects the trajectory Δ_1 with a slope equal to or greater than $-K$. The above intersection point is denoted by R in Figure 3(a). For the case in which the point P is on the trajectory Δ_1, the points P and R coincide with each other. Since the slope of the trajectory Δ_1 is K, the slope of the trajectory Γ at the point P is less than K. By the above arguments the functions $\phi(T)$ and $\psi(T)$ satisfy Relations (16) and (17) at the point R. Therefore, the controlled quantity Y can be kept identically zero from the instant designated by T_R which corresponds to the point R as shown in Figure 3(a).

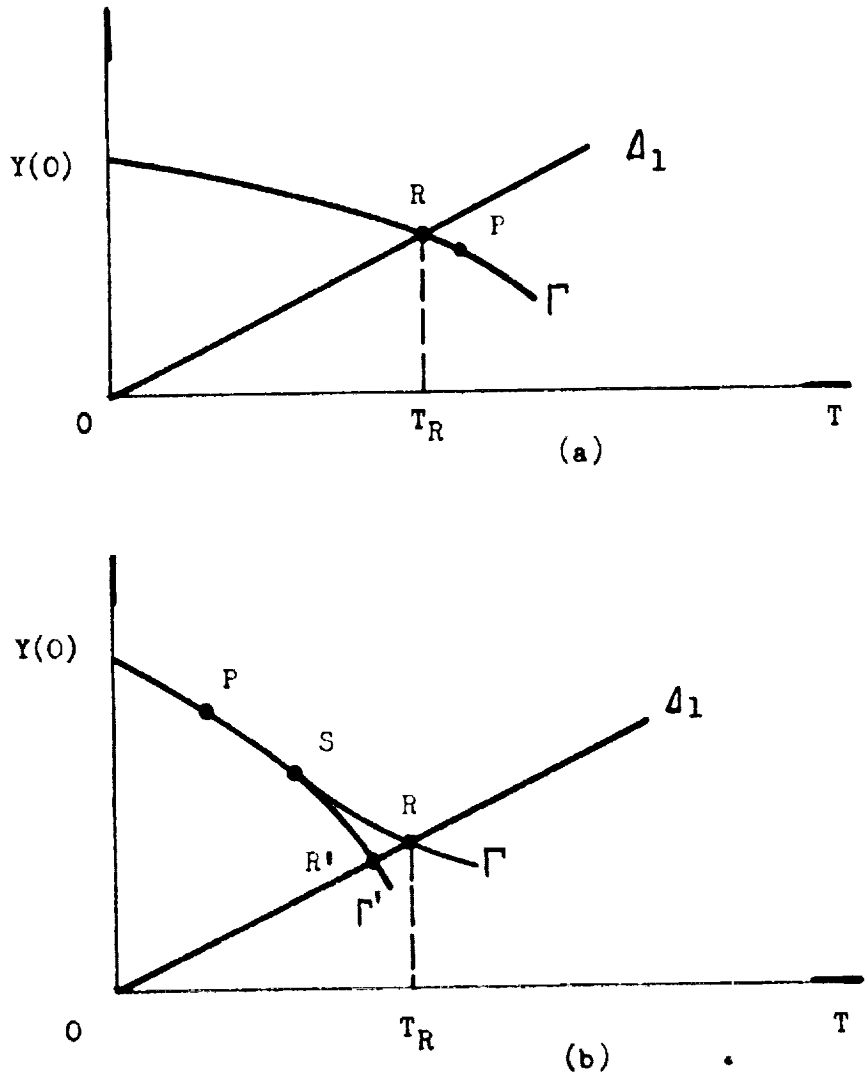


FIGURE 3 TRANSIENT TYPES C

(a) TRANSIENT TYPE C_1

(b) TRANSIENT TYPE C_2

For the second case, the trajectory determined by the plot of Relation (27) has slope -K at a point above the trajectory Δ_1. This case is illustrated in Figure 3(b). Denote the curve APR' which corresponds to Relation (27) by Γ'. The point at which the trajectory Γ' has slope -K is designated by P. The point P may exist for $T \leq 0$. The point R' is the intersection point of the trajectories Γ' and Δ_1. Since the trajectory Γ' is concave down, it has a slope less than -K at the point R'. Since Relation (16) is not satisfied at the point R', the controlled quantity Y cannot be made identically zero. The optimum Φ trajectory will be obtained, then, by the following procedure. Choose a point on the curve segment PR', calling this point the switching point S. At the point S we set $\Phi'' = 1$ so that the concavity of the Φ trajectory is up and maximum. We adjust the point S such that the Φ trajectory intersects the trajectory Δ_1 with slope -K. The Φ trajectory obtained above is designated by Γ in Figure 3(b). The intersection point of the trajectories Γ and Δ_1 is denoted by R.

That such points S and R exist may be seen as follows. Let the point S coincide with the point P. The trajectory Γ' has a slope greater than -K so that the slope of the trajectory Γ at the point R is also greater than -K. As we move the point S along the trajectory Γ' for increasing time, the slope of the trajectory Γ at the point R continuously decreases and the point R approaches the point R'. When the point S reaches R', the points S, R' and R coincide together and the slope of the trajectory Γ at this point is less than -K. As shown above the slope of the trajectory Γ at the point R changes from a value greater than -K to a value less than -K as the point S moves from P to

R'. Hence, there must exist a point S on the trajectory Γ' such that the trajectory Γ has s slope -K at the point R.

The functions $\Phi(T)$ and $\Psi(T)$ determined by Γ and Δ_1 of Figure 3(b) satisfy Relations (16) and (17) at the instant designated by T_R which corresponds to the point R. Hence, the controlled quantity Y can be kept identically zero from the instant $T = T_R$. It will be proved in the next section that the Φ and Ψ trajectories determined in the above two cases are optimal. The type of optimum transient for the first case is called Transient Type C_1 and for the second case Transient Type C_2. Transient Types C_1 and C_2 are the class of transient types denoted by C. As shown above, for Transient Types C, the initial selection of the controlling variables is $X_1' = -1$ and $X_2 = K$.

For the next type of transient called Transient Type D, we treat the trajectory designated by Γ_3 in Figure 2 which corresponds to Relation (24). The initial slope of the trajectory Γ_3 is between the initial slopes of the trajectories Γ_2 and Γ_4 in magnitude. Therefore, the point denoted by K_3 at which the trajectory Γ_3 has slope -K is between the trajectories Δ_1 and Δ_2. The optimal trajectories representing the functions $\Phi(T)$ and $\Psi(T)$ are determined as follows. Plot the trajectories Γ_3, Δ_1 and Δ_2 again in Figure 4. Draw the tangential line to the trajectory Γ_3 at the point P_3, i.e. let the tangent line have slope -K. Let the intersection point of the above tangent line and the trajectory Δ_1 be denoted by S which is called the switching point. Let the line composed of the line segments OS and SP_3 be designated by Δ. The line segment SP_3 corresponds to the equation $\Phi' = -K$. Relations (16) and (17) are satisfied by the functions $\Phi(T)$

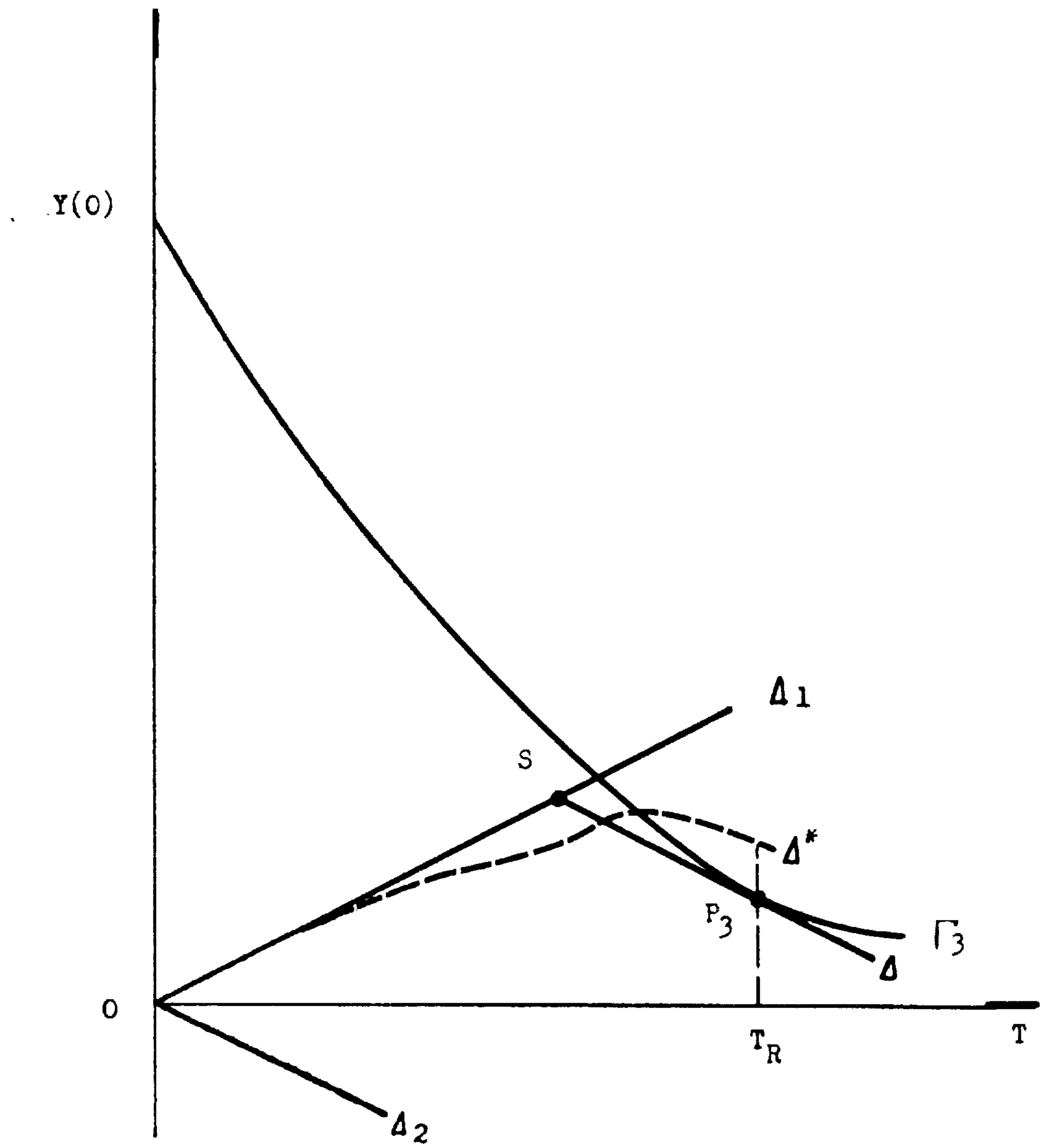
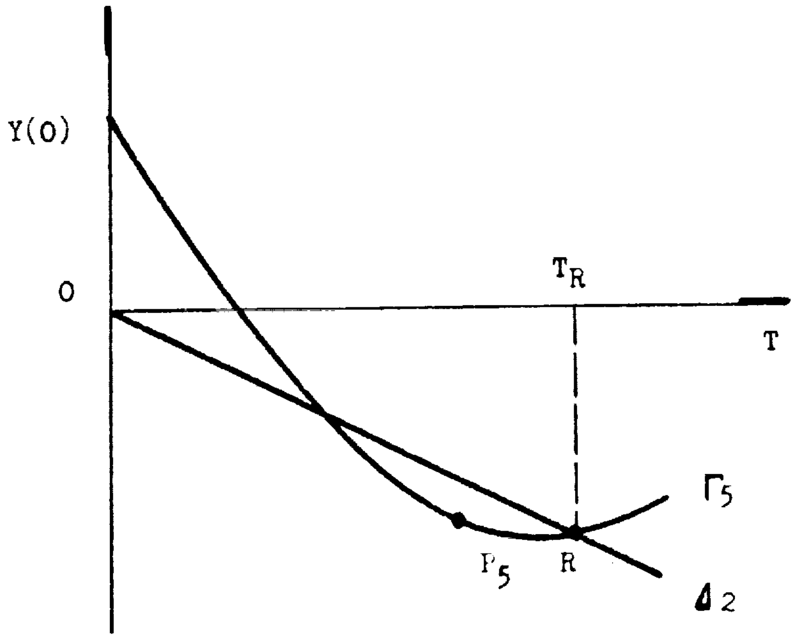


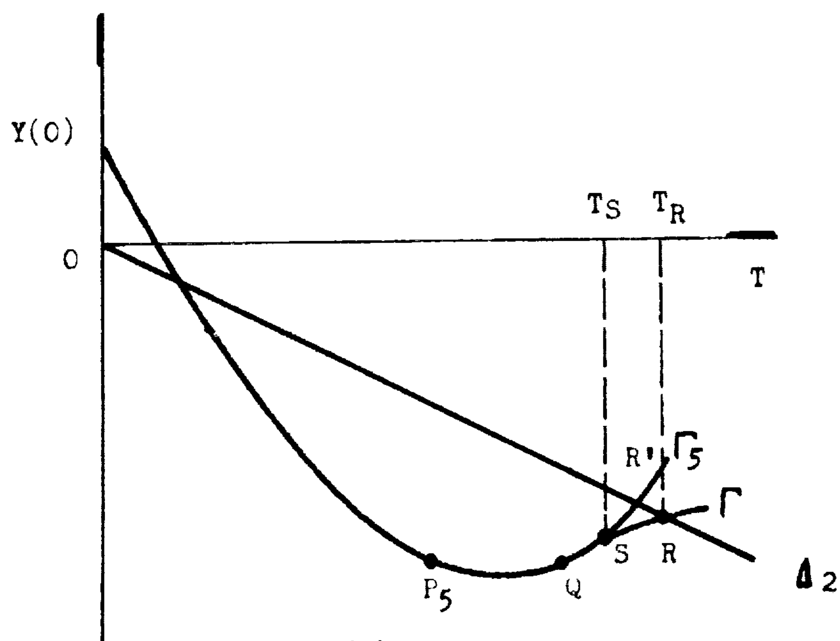
FIGURE 4 OPTIMUM TRAJECTORIES FOR TRANSIENT TYPE D

and $\Psi(T)$ representing the trajectories Γ_3 and Δ_1, respectively, at the instant $T = T_R$ where the time T_R corresponds to the point P_3. Hence, the controlled quantity Y can be made identically zero for $T \geq T_R$. The proof that the trajectories Γ_3 and Δ yield the optimum transient for Stage 1 will be given in the next section. Thus, it has been shown that the initial setting of the controlling variables for Transient Type D must be $X_1' = 1$ and $X_2 = -K$ for optimum control.

Now consider the plot of Relation (24) represented by the trajectory Γ_5 in Figure 2 where the point P_5 on the trajectory Γ_5 lies below the trajectory Δ_2. Again P_5 denotes the point on the trajectory Γ_5 where the slope is -K. This case may occur when the value of $X_1(0)$, i.e. $\Phi'(0)$ is small. Since the trajectory Γ_5 is concave up, it intersects the trajectory Δ_2 twice. At the first intersection point the slope of the trajectory Γ_5 is less than -K so that Relation (16) is not satisfied, i.e. the controlled quantity Y cannot be made identically zero. The slope of the trajectory Γ_5 is always increasing at the maximum rate and becomes -K at the point P_5, where Relation (16) is satisfied for the first time. However, Relation (17) is not satisfied at the point P_5 so that again the controlled Y cannot be made zero. After the point P_5 is attained, the trajectory Γ_5 intersects again the trajectory Δ_2. Two cases are considered with respect to the value of the slope of the trajectory Γ_5 at the second intersection point of the trajectories Γ_5 and Δ_2. For the first case the slope of the trajectory Γ_5 at the second intersection point is equal to or less than K. The trajectories Γ_5 and Δ_2 for this case are shown in Figure 5(a). The second intersection of the two trajectories is denoted by R,



(a)



(b)

FIGURE 5 OPTIMUM TRAJECTORIES FOR TRANSIENT TYPES E and the corresponding time by T_R. The trajectory Γ_5 has a slope greater than -K at the point R. By the above arguments the functions $\Phi(T)$ and $\Psi(T)$ corresponding to the trajectories Γ_5 and Δ_2, respectively, satisfy Relations (16) and (17) at the instant $T = T_R$. Hence, the controlled quantity Y can be made identically zero for $T \geq T_R$.

Consider the second case where the trajectory Γ_5 has a slope greater than K at its second intersection point with the trajectory Δ_2 which is designated by R' as shown in Figure 5(b). The trajectory Γ_5 must have the slope -K at some point on the curve segment of the trajectory Γ_5 determined by the points P_5 and R'. If the concavity of the trajectory Γ_5 is changed at some point denoted by S on the curve segment PR' by setting $\phi'' = -1$, we obtain the trajectory designated by Γ, where Γ from the point S on is concave down with its slope decreasing at the maximum rate. By reasoning similar to that employed for Transient Type C_2, the point S can be chosen so that the trajectory Γ will cross the trajectory Δ_2 with the slope K. This may be understood if we consider the mirror images, with respect to the T axis, of the trajectories Γ, Γ' and Δ_1 in Figure 3(b) and compare them with the trajectories Γ, Γ_5 and Δ_2 in Figure 5. Let the intersection point of the trajectories Γ and Δ_2 be denoted by R and the corresponding instant by T_R. The trajectories Γ and Δ_2 determine the functions $\Phi(T)$ and $\Psi(T)$ which satisfy Relations (16) and (17) at the instant $T = T_R$ so that the controlled quantity Y can be kept identically zero from that instant. The type of transient for the first case is called Transient Type E_1, and for the second case Transient Type E_2. Transient Types

E_1 and E_2 are classed as Transient Types E. It will be proved in the next section that Transient Types E are optimum. As shown above, for Transient Types E the controlling variables at least near $T = 0$ must be set $X_1' = 1$ and $X_2 = K$.

We have treated above all possible cases of optimum transients occuring for the initial conditions $Y(0) \neq 0$. Now we are going to deal with the optimum control for the case in which $Y(0) = 0$. As explained in the begining of this section it is sufficient to treat only the case in which $X_1(0) > 0$, in investigating the optimum transients with the initial condition $Y(0) = 0$. This type of transient is called Transient Type F. If $Y(0) = 0$ and $0 < X_1(0) \leq K$ Relations (16) and (17) are satisfied. Hence, the controlled quantity Y can be made identically zero for $T \geq 0$, i.e. Stage 1 does not exist. For $Y(0) = 0$ and $X_1(0) > K$, Transient Type F is treated below. For $T = \varepsilon > 0$, Y becomes positive, for any X_1' and X_2 at $T = 0$ under Conditions (7) and (8), for some positive number ε. The reason for this is shown below. Since $X_1(0) > K$ by the assumption made for Transient Type F and $|X_2| \leq K$ by Relation (9) it follows that

$$Y' = X_1 + X_2 > 0 \qquad \text{at } T = 0 \tag{28}$$

Hence, we have

$$Y(\varepsilon) = \int_0^{\varepsilon} Y' \, dT + Y(0) > 0 \tag{29}$$

If we choose the time origin at $T = \varepsilon$ we obtain the initial condition $Y(0) > 0$. Therefore, Transient Type F reduces to Transient Type C.

PROOF THAT TRANSIENTS ARE OPTIMUM

In the last section we treated all possible transient types which may occur in Stage 1. We shall prove in this section that these types of transients are all optimum and uniquely determined so that the quantity X_1' and X_2 must be kept always at their extremum values in Stage 1. It was shown that Transient Types A and B are optimum. It was also shown that Transient Type C_2 is included by Transient Type E_2. By the same reason Transient Type C_1 is involved by Transient Type F_1. It was explained that the treatment of Transient Type F is equivalent to that for Transient Type C. Thus, in this section, it is sufficient to prove that Transient Types D, E_1 and E_2 are optimum.

Before discussing the problem further we define the controlled area as follows. The controlled area is that area between a Φ trajectory and a Ψ trajectory in Stage 1.

The optimality of Transient Type D is proved first. We shall prove that the trajectories Γ_3 and Δ in Figure 4 determine the unique optimum transient. As explained in the last section a transient time T_R is obtained such that the trajectory Γ_3 with slope -K coincides with the trajectory Δ at $T = T_R$. Since the slope of Γ_3 is increasing at the maximum rate and is equal to -K at $T = T_R$, the slope of Γ_3 is always less than or equal to -K for $T < T_R$. Also, any Φ trajectory other than Γ_3 cannot have a slope equal to or greater than -K for $T \leq T_R$. Hence the time duration T_R is the minimum transient time and the trajectory

Γ_3 is optimum and uniquely determined. We will prove that the trajectory Δ yields the minimum controlled area while the minimum transient time T_R is kept. The proof is by contradiction. Assume that there exists a Ψ trajectory called Δ^* which yields less controlled area than the trajectory Δ and coincides with the trajectory Γ_3 at the instant $T = T_R$. Since the trajectory Δ lies under Γ_3 for $T < T_R$, the trajectory Δ^* must be over some finite time interval above Δ in order to make the controlled area less. However, the trajectory Δ^* cannot lie above Δ since Δ is the upper boundary of all trajectories Δ^* for the time interval corresponding to the segment OS. Therefore, the trajectory Δ^* must exist, at least partly, above the line segment OS. However, the trajectory Δ^* having passed the time corresponding to the point S and once exceeding the line segment SP cannot reach this segment again. The reason for this is that the trajectory Δ is decreasing at the maximum rate for the line segment SP. Therefore, the trajectory Δ^* cannot coincide with Γ_3 at the instant $T = T_R$. This violates the assumption previously stated for the trajectory Δ^*. Hence, there exists no Ψ trajectory yielding less controlled area than Δ. As seen from Figure 4, the controlled quantity Y for Transient Type D is always positive and decreasing monotonically, i.e. the transient has no over or under swing. This concludes the proof of optimality of Transient Type D.

Next, we treat Transient Type E_1. We shall prove below that the trajectories Γ_5 and Δ_2 in Figure 5(a) determine the optimum transient. These trajectories are plotted in Figure 6 with notation unchanged. Consider any trajectories for $\Phi(T)$ and $\Psi(T)$ where the Φ trajectory has the same initial conditions as those for Γ_5. We plot the above trajectories of Φ and Ψ, and designate them as Γ_5^* and Δ_2^*,

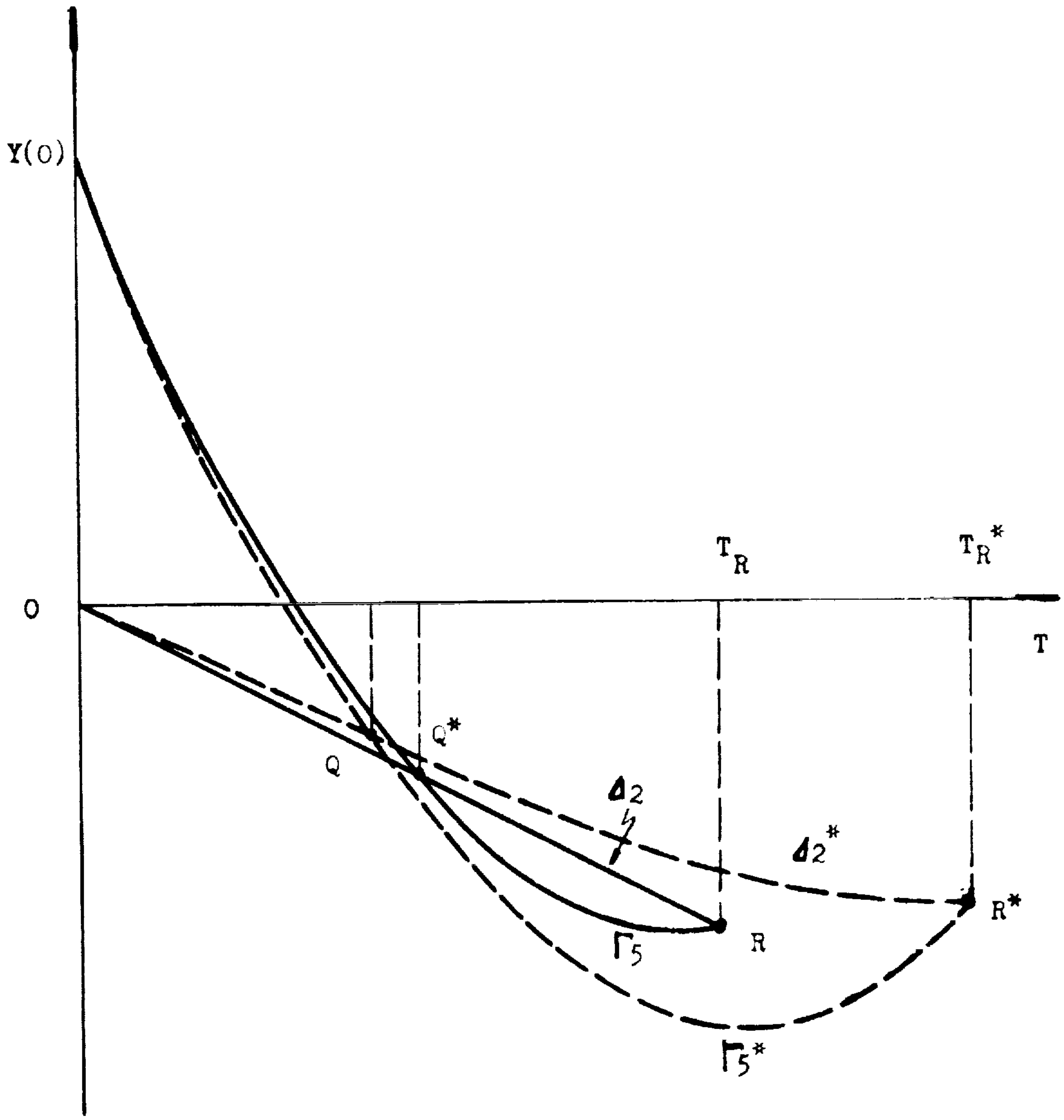


FIGURE 6 OPTIMUM AND NON-OPTIMUM TRAJECTORIES FOR TRANSIENT TYPE E_1

respectively, in Figure 6. Here, the trajectory Δ_2 is the lower boundary of Δ trajectories. Since the trajectory Γ_5 is concave up with its slope increasing at the maximum rate, the trajectory Δ_2^* exists below Δ_2. Suppose that the trajectories Γ_5^* and Δ_2^* intersect each other twice. The case in which the second intersection point does not exist will be treated in the same manner as for the above case, i.e. consider that the second intersection point exists at $T = \infty$. Let the first intersection point be denoted by Q^* and the second by R^*. Also let the points at which the trajectories Γ_5 and Δ_2 intersect be designated by Q and R as in Figure 6. Since the trajectories Δ_2^* and Γ_5^* lie above and below, respectively, the trajectories Δ_2 and Γ_5, the point Q^* exists before the point Q and R^* after R as shown in Figure 6. We will show first that the controlled quantity Y cannot be made identically zero at the instant corresponding to the point Q^*. Since the slope of the trajectory Γ_5 is increasing at the maximum rate and becomes -K at the point P_5, the trajectory Γ_5^* cannot have a slope greater than or equal to -K before the time corresponding to the point P_5. Hence, the slope of Γ_5^* at the point Q^* is less than -K and Relation (17) is not satisfied. Thus the quantity Y cannot be made identically zero. Since the second intersection point R^* exists after the time $T = T_R$ the transient time determined by the trajectories Γ_5^* and Δ_2^* cannot exist before the instant $T = T_R$. This is sufficient to show that the trajectories Γ_5^* and Δ_2^* are not optimum and the unique optimum trajectories are determined by the trajectories Γ_5 and Δ_2.

Finally, we treat Transient Type E_2. The trajectories Γ and Δ_2 shown in Figure 5(b) are plotted again in Figure 7 with notation unchanged. We shall prove below that the time duration T_R is the minimum transient time and is uniquely determined by the trajectories Γ and Δ_2. Consider another pair of Φ and Ψ trajectories and designate them respectively as Γ^* and Δ_2^*. Since the trajectory Δ_2 is the lower boundary of Ψ trajectories, Δ_2^* may not be below Δ_2. The trajectory Γ is concave up with its slope increasing at the maximum rate for $0 \leq T < T_S$ so that the trajectory Γ^* cannot lie above Γ for this time interval. The trajectories Γ^* and Δ_2^* intersect each other at a time $T \leq T_S$, but the controlled quantity Y cannot be set identically zero at that instant. The reason for this is the same as that for Transient Type E_1 shown already.

The point on Γ^* at the time $T = T_S$ denoted by S^* in Figure 7 can exist only below the point S. Let T_R^* be the transient time when the trajectories Γ^* and Δ_2^* intersect. Suppose that the time T_R^* is less than T_R. Then the trajectory Γ^* must cross the trajectory Γ with its slope greater than that of Γ at some time T between T_S and T_R^*. But the slope of Γ is decreasing at the maximum rate. Hence, the slope of Γ^* is greater than that of Γ at the instant $T = T_R^*$. Also, the trajectory Γ has greater slope at the time $T = T_R^*$ than at $T = T_R$. Therefore, the slope of Γ^* at the instant $T = T_R^*$ is greater than K. Thus, Relation (16) does not hold for the function $\Phi(T)$ represented by the trajectory Γ^* at the instant $T = T_R^*$, i.e. the trajectories Γ^* and Δ_2^* are not admissible. It was proved above that the minimum transient time must be T_R and this is uniquely determined by the trajectories Γ and Δ_2. As seen from Figures 6 and 7, non optimum trajectories yield greater under swing than the optimum trajectories. This concludes the proof of optimality for Transient Type E_2.

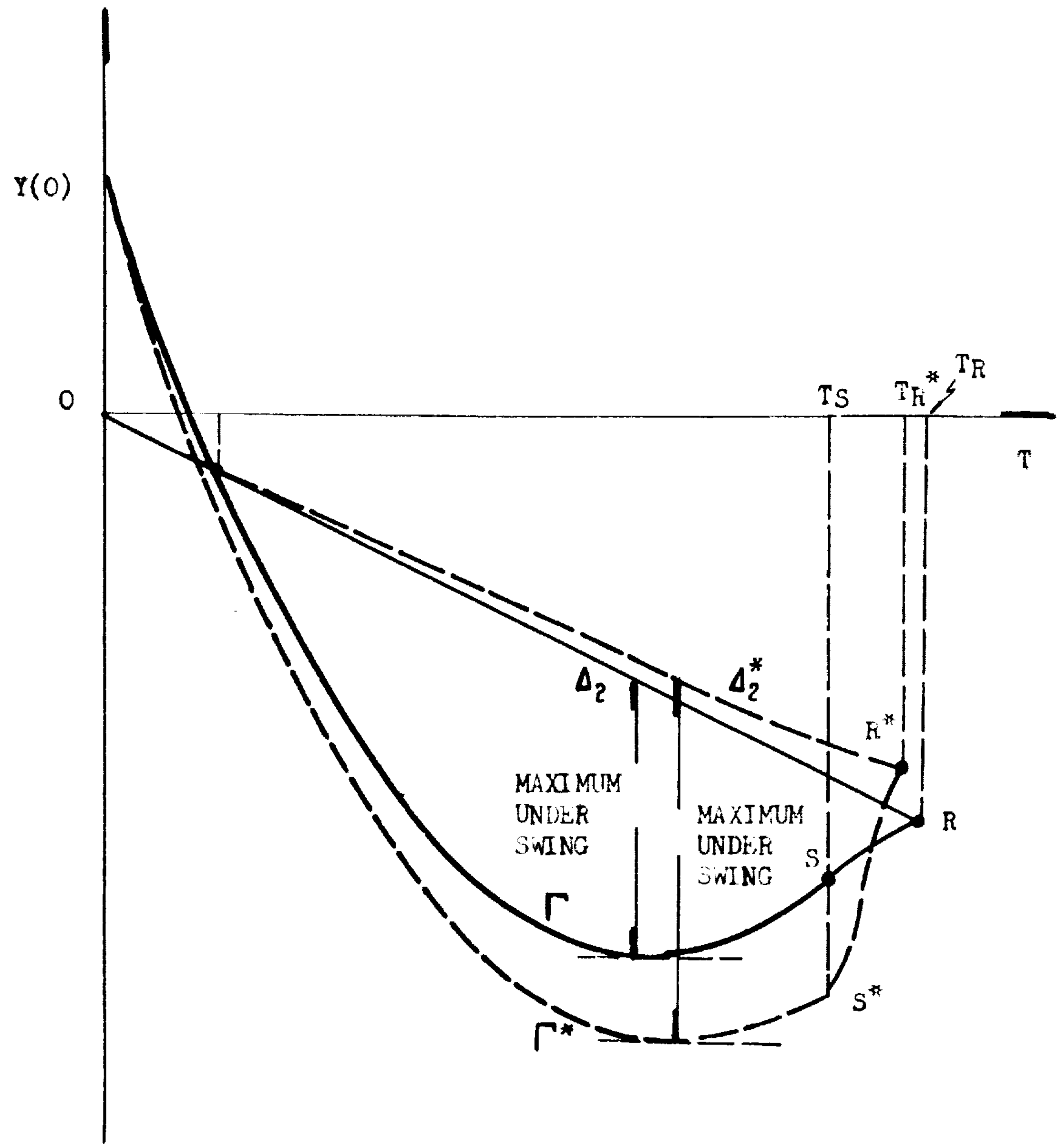


FIGURE 7 OPTIMUM AND NON-OPTIMUM TRAJECTORIES FOR TRANSIENT TYPE E_2

CONTROL FUNCTIONS

In the previous sections all the possible optimum transients in Stage 1 were shown and proofs of optimality were given. A control which always gives the optimum transient must be able to predict which type of optimum transient will follow so that the proper values of X_1' and X_2 may be chosen. The functions, called the first and second control functions, will be introduced in this section. The sign of each of these control functions which depends on the initial conditions is used to determine the optimum selection of X_1' and X_2.

The first control function is used in distinguishing whether trajectory obtained by setting $X_1' = 1$ is above, on or below the Φ trajectory Γ_2 shown in Figure 2. When the Φ trajectory is below Δ_2, the second control function is used to determine whether the Φ trajectory is above, on or below the trajectory Γ_4.

Derivation of First Control Function

Assume that $Y(0) > 0$. We consider what relationship between $Y(0)$ and $X_1(0)$ will exist when the trajectory Γ_2 as shown in Figure 2 is obtained. Since the slope of the trajectory Γ_2 at the time T_{R2} corresponding to the point P_2 is $-K$, it follows from Relation (24) that

$$\Phi' = T_{R2} + X_1(0) = -K \qquad (30)$$

The time T_{R2} given by the above relation is negative or zero when

$$X_1(0) \geq -K \qquad (31)$$

In this case the optimal values of X_1' and X_2 are made independent of the control functions. This is discussed in the Appendix. Hence, we assume here that

$$X_1(0) < -K \tag{32}$$

At the time $T = T_{R2}$ the trajectories Γ_2 and Δ_1 coincide with each other so that $\Phi = \Psi$. Therefore, by Relations (21) and (24) we have the following expression

$$\Phi - \Psi = \frac{T_{R2}}{2} + X_1(0)\,T_{R2} + Y(0) - KT = 0 \tag{33}$$

where the positive sign in the right hand side of Relation (21) has been chosen. Substituting T_{R2} of Relation (30) into Relation (33), and arranging it we have

$$\Phi - \Psi = Y(0) + \tfrac{1}{2}\{K + X_1(0)\}\{3K - X_1(0)\} = 0 \tag{34}$$

Next we assume that $Y(0) < 0$ which corresponds to the mirror image about the T axis of the trajectory Γ_2 shown in Figure 2. The plot of the above mirror image of the trajectory Γ_2 is designated by Γ_2^* in Figure 8 while the mirror image of the trajectory Δ_1 is the trajectory Δ_2. The point P_2^* is the mirror image of the point P_2, of Figure 2. Here the trajectory Γ_2^* intersects the trajectory Δ_2 with a slope K at the point P_2^*. Following the derivation of Relation (34), we can find the expression at the time T_{R2}^* which corresponds to the point P^*, i.e.

$$\Psi - \Phi = -Y(0) + \tfrac{1}{2}\{K - X_1(0)\}\{3K + X_1(0)\} = 0 \tag{35}$$

We assume, as in the first case that the time T_{R2}^* is positive, i.e. referring to Relation (26) we have

$$X_1(0) > K \tag{36}$$

The case for which $X_1(0) \leq K$ is referred to the Appendix.

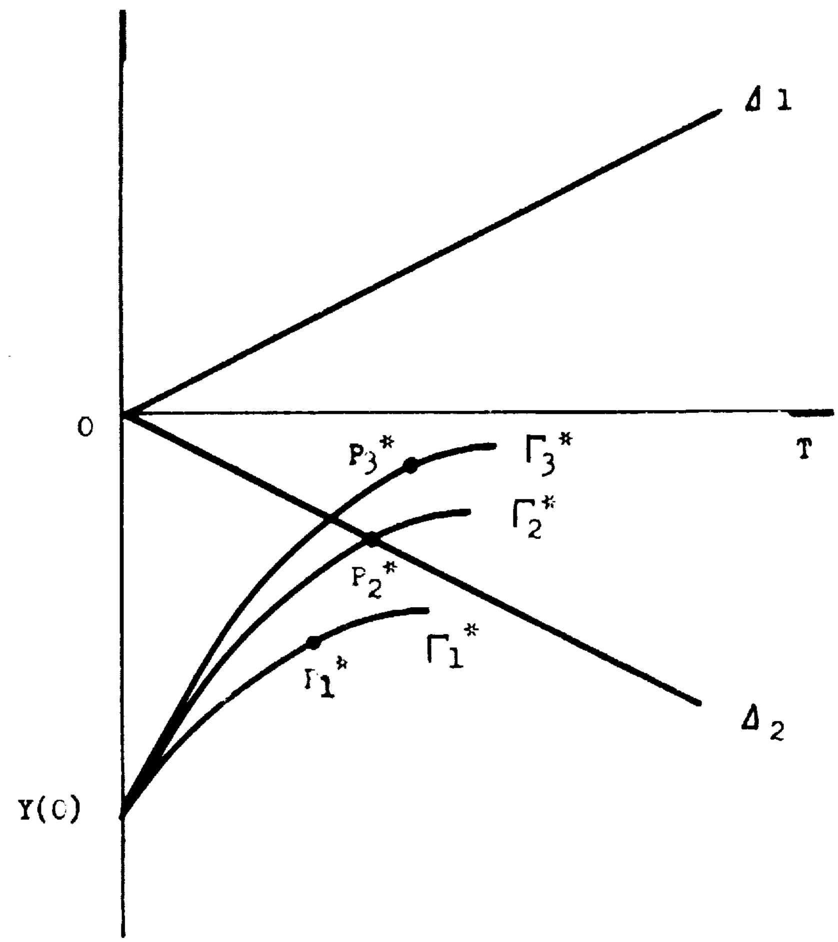


FIGURE 8 MIRROR IMAGES OF TRAJECTORIES Γ_1 , Γ_2 AND Γ_3

Equations (34) and (35) may be combined into the single equation

$$\Sigma_1 = 0 \tag{37}$$

where Σ_1, called the first control function, is given by

$$\Sigma_1 = |Y| + \tfrac{1}{2}(K - |X_1|)(3K + |X_1|) \tag{38}$$

provided

$$Y > 0, \quad X_1 < -K \tag{39}$$

or

$$Y < 0, \quad X_1 > K \tag{40}$$

Here $Y(0)$ and $X_1(0)$ have been replaced by Y and X_1 respectively since the time axis may be shifted arbitrarily.

As may be seen from the definition of the first control function, it has the following properties. The case where $\Sigma_1(0) > 0$ is represented by the trajectory Γ_1 of Figure 2 or Γ_1^* of Figure 8. Here the trajectory Γ_1^* is the mirror image of Γ_1. Similarly the case in which $\Sigma_1(0) < 0$ corresponds to the trajectory Γ_2 of Figure 2 and Γ_3^* of Figure 8, the trajectories Γ_3 and Γ_3^* are in mirror image relationship.

Derivation of Second Control Function

We are going to derive the relation in terms of the quantities $X_1(0)$ and $Y(0)$ corresponding to the trajectory Γ_4 shown in Figure 2. Thus assume first $Y(0) > 0$. The relation in consideration may be derived in the same manner as that for derivation of Relation (34) except that we choose here the negative sign in the right hand side of Relation (21). After calculation it follows that

$$\Phi - \Psi = Y(0) - \tfrac{1}{2}\{K + X_1(0)\}^2 = 0 \tag{41}$$

Assuming next that $Y(0) < 0$, we can derive the relation corresponding to Relation (35), i.e.

$$\psi - \Phi = -Y(0) - \tfrac{1}{2}\{K - X_1(0)\}^2 = 0 \qquad (42)$$

Similarly as in the previous section Relations (41) and (42) may be combined into the single equation

$$\Sigma_2 = 0 \qquad (43)$$

where Σ_2 is called the second control function and is given by

$$\Sigma_2 = |Y| - \tfrac{1}{2}(K - |X_1|)^2 \qquad (44)$$

Here again $X_1(0)$ and $Y(0)$ have been replaced by X_1 and Y respectively. In discussing geometrical properties of the second control function, the argument for the first control function holds here after replacing the trajectories Γ_2 and Δ_1 respectively by Γ_4 and Δ_2.

As seen from the derivation of the first and second control functions, if $\Sigma_1 \geq 0$, it always follows that $\Sigma_2 > 0$. The only case when the sign of Σ_2 cannot be determined by observing the sign of Σ_1 is when $\Sigma_1 < 0$. Hence, as seen in the next section, we design the control schedule such that the sign of the first control function is checked first. If this sign is negative, the sign of the second control function is monitored.

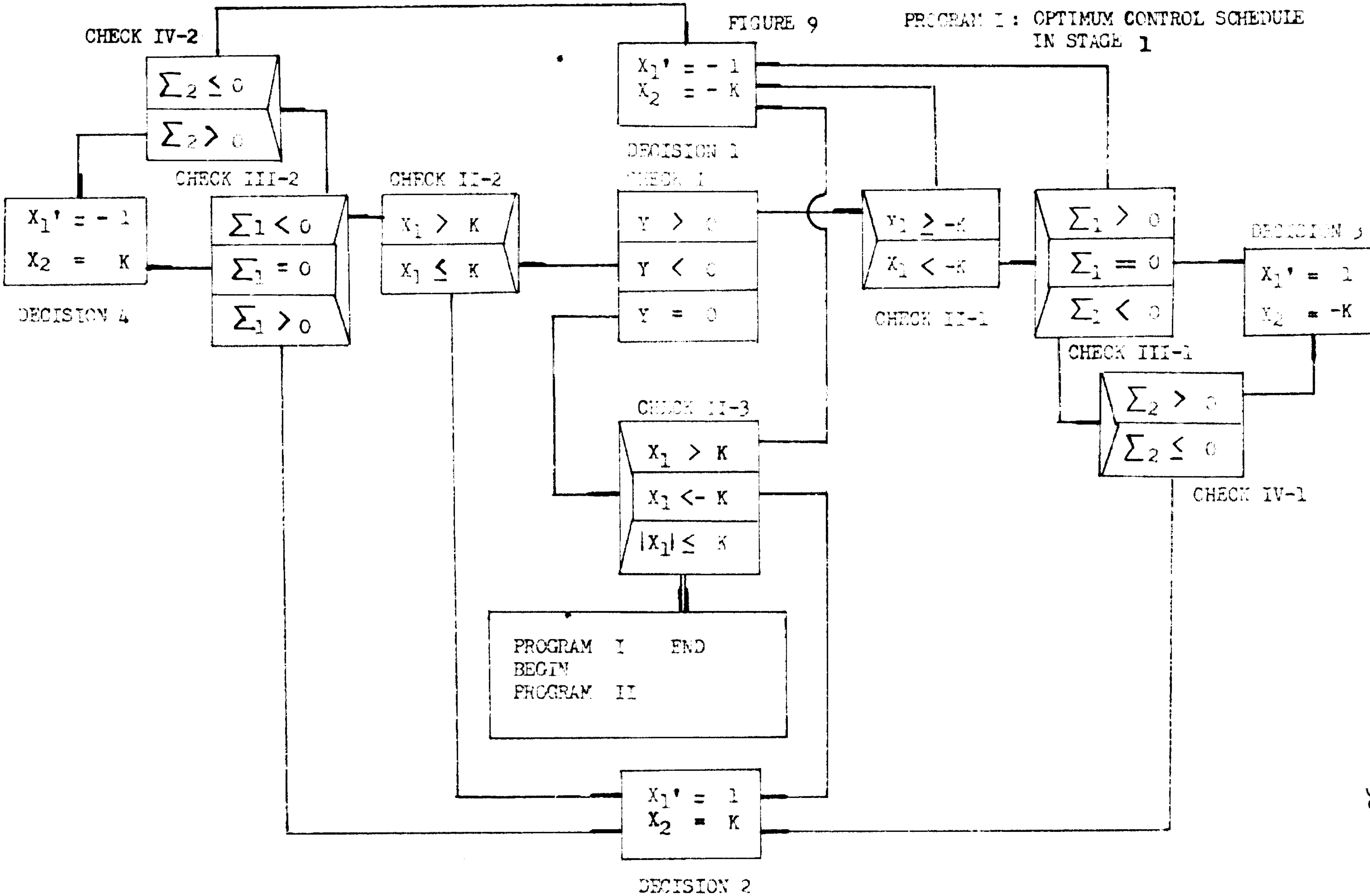
OPTIMUM CONTROL SCHEDULE IN STAGE 1

With the aid of the control functions obtained in the last section we can obtain all the information needed for determining the optimal values of X_1' and X_2. A schematic diagram of an optimum schedule, called Program I, is shown in Figure 9. This schedule always gives the optimal choice of X_1' and X_2 in Stage 1 based on the instantaneous values of $Y(T)$ and $X_1(T)$. Hence, Program I yields one of Transient Types A through F. Interpretation of the schematic diagram shown in Figure 9 will be given below with an example.

Program I is composed of two kind of elements, called Checks and Decisions. The function of a Check is to monitor a state of the system at every instant and the function of a Decision is to determine the optimal selection of X_1' and X_2 based on Checks. A Check and Decision process always begins with Check I and ends at one of Decisions 1 through 5. A series of Check and Decision processes will continue until Decision 5 is reached. If Decision 5 is attained, Program I stops its function and Program II which is explained in the next section begins.

Consider, for example, the case in which Transient Type E_2 as shown in Figure 5(b) takes place. Since $Y(0)>0$, monitoring of the state of the system goes from Check I to Check II-1. Sensing $X_1(0) < -K$, Program I monitors next Check III-1, where $\Sigma_1(C) < 0$ is obtained. Hence, Check IV-1 must be made and $\Sigma_2(0) < 0$ is sensed. This results in Decision 2, where $X_1' = 1$ and $X_2 = K$ are selected. This selection agrees with the

FIGURE 9 PROGRAM I : OPTIMUM CONTROL SCHEDULE IN STAGE 1



result explained in Trasient Type E_2. The above Check and Decision process does not change until the first intersection point of the trajectories Γ and Δ_2 is obtained. At the first intersection point the state $Y = 0$ is sensed by Check I. Then Check II-3 must be made where $X_1 < -K$ is monitored so that Decision 2, the same as before, is selected. Instantaneously, the trajectory Γ goes below Δ_2 and $Y < 0$ holds in Check I. Therefore, Check II-2 must follow and Decision 2 results. The above process continues until the point P on Γ is obtained where $Y < 0$ and $X_1 = K$. At the instant after the point P is obtained there occurs a change in Check II-2, i.e. $X_1 > K$ instead of $X_1 \leq K$ holds. However, since $\Sigma_1 > 0$ at this instant, Decision 2 is still kept by the result of Check III-2. At $T = T_S$, it follows that $Y < 0$, $X_1 > K$ and $\Sigma_1 = 0$. For the first time Decision 2 is cancelled at this instant and Decision 4 is chosen through Check I, Check II-2 and Check III-2. This process continues until the time $T = T_R$. At $T = T_R$ we obtain $Y = 0$ and $X_1 = K$ and Decision 5 is selected. Program I ends and Stage 1 terminates at this instant.

OPTIMUM CONTROL SCHEDULE IN STAGE 2

Previous sections concerned the optimum control in Stage 1. The controlled quantity is brought to zero at the end of Stage 1. This does not imply, however, that the controlling variables X_1 and X_2 are also brought to zero. In this section we shall develope a control schedule in which the variables X_1 and X_2 become zero in minimum time while the controlled quantity Y is kept zero.

When Stage 1 terminates it follows that

$$Y = 0 \tag{45}$$

$$|X_1| \leq K \tag{46}$$

We take the time origin at the end of Stage 1. Then the above relations are the initial conditions in Stage 2. In order to keep Y identically zero throughout Stage 2 we must set the derivative Y' also zero, i.e.

$$Y' = X_1 + X_2 = 0$$

or

$$X_2 = - X_1 \tag{47}$$

Since the variable X_2 can be made to take on any value subject to Restriction (8) and since the variable X_1 has the initial condition as in Relation (46), the variables X_1 and X_2 satisfy Relation (47) at T = 0.

The optimum control in Stage 2 is developed below. The variables X_1 and X_2 must approach zero in minimum time in such a manner that Relation (47) is satisfied. The variable X_2 can be set zero at any time but only the derivative of X_1 may be controlled. Hence, the optimum control is obtained when the variable X_1 approaches zero at the maximum rate. This implies the following: if the variable X_1 is positive its derivative X_1' must be -1. Similarly, if the variable X_1 is negative the derivative X_1' must be 1. In both cases the absolute value of X_1 decreases toward zero so that Relation (46) holds not only at T = 0 but also at any instant in Stage 2. Whenever Relation (46) holds Relation (47) can be satisfied as previously explained. If the variable X_1 is zero, the variable X_2 is also set zero.

The optimum control schedule developed above is shown by a schematic diagram called Program II as in Figure 10. The interpretation of this diagram is as follows. Program II begins if and only if Program I terminates. As in Program I, Program II has two kind of elements, Check and Decision. An optimum process must go to Check first where the sign of X_1 is monitored. If the variable X_1 is positive, Decision 1 will be selected and $X_1' = -1$ and $X_2 = - X_1$. A similar argument applies to the case when the variable X_1 is negative. When both X_1 and X_2 vanish Decision 3 is selected and the system remains at the reference state. At this time Stages 1 and 2 terminate.

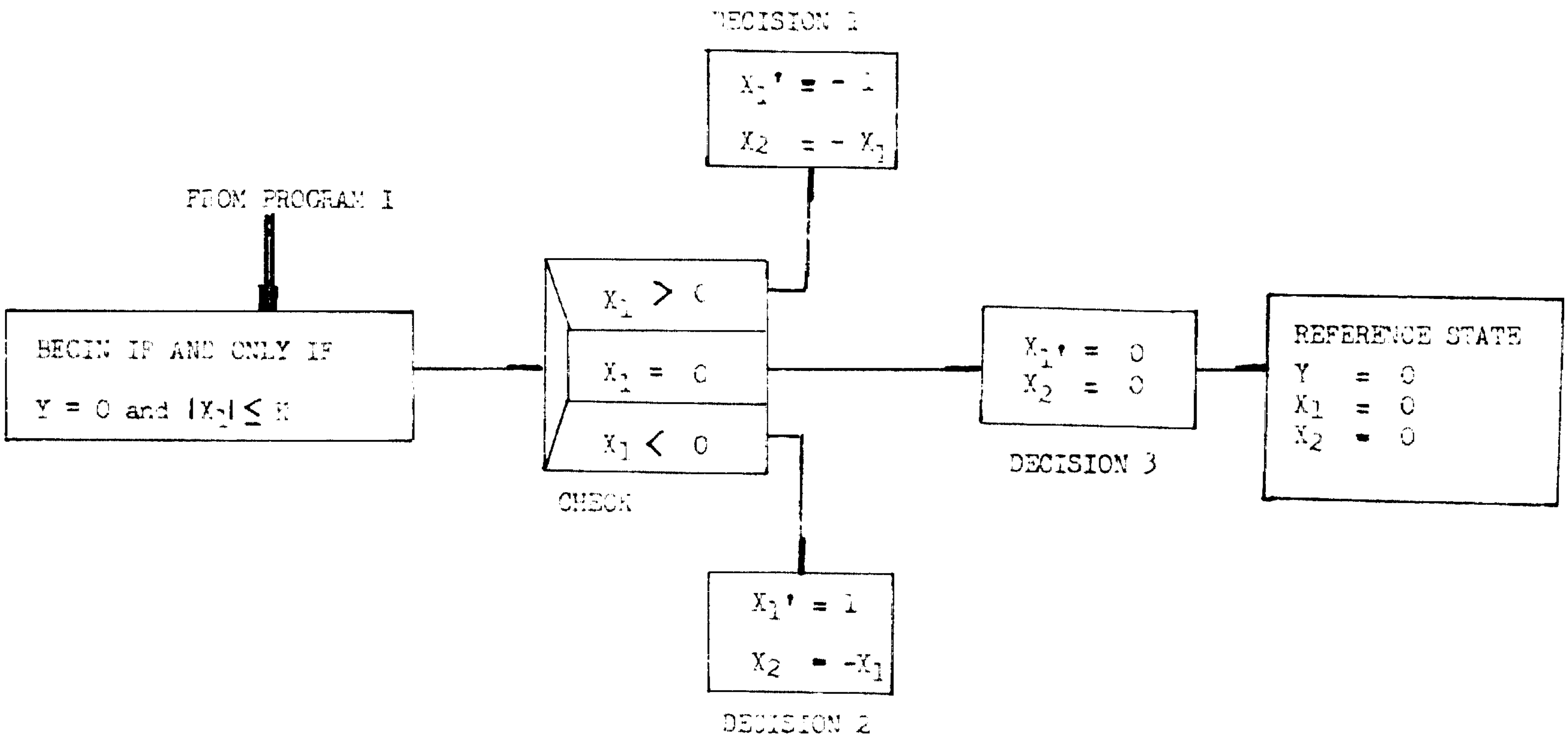


FIGURE 10 PROGRAM II : OPTIMUM CONTROL SCHEDULE IN STAGE 2.

ANALOG COMPUTER SIMULATION RESULTS

The results of the optimum control developed in the previous sections were verified on an electronic analog computer. Two examples are illustrated in Figures 11 (a) and (b). Here an upper figure shows the transient of the controlled quantity Y for each case and a lower figure refers to the corresponding response of the controlling variables X_1 and X_2. Figures 11 (a) show Transient Type F and Figures 11 (b) Transient Type E_1.

Relays were used in the logical switching circuits both in Programs I and II. As seen in lower figures, the controlling variables fluctuate momentarily at the relays switching instance. However, this does not significantly affect the controlled quantity since the controlled quantity is the integrated value of the controlling variables and integration has a low pass filter effect which smoothes the response. In fact, as shown in the upper figures of Figures 11, the actual responses of the controlled quantity deviate from the theoretical values only near T = 0 and at the end of Stage 1.

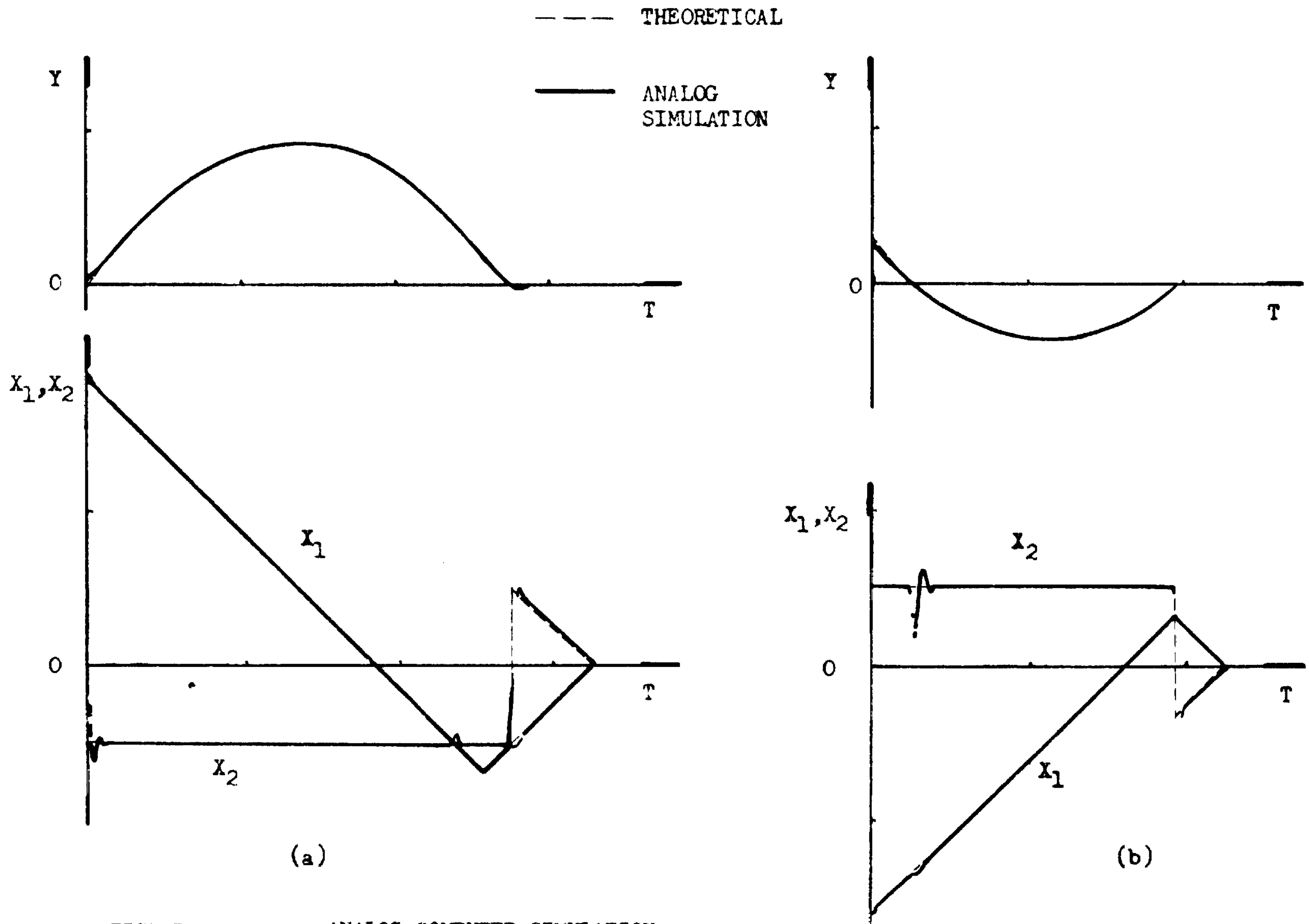


FIGURE 11 ANALOG COMPUTER SIMULATION OF OPTIMUM CONTROL

CONCLUSION

A direct and straightforward technique to solve a simple class of optimum nonlinear multivariable control probelms has been presented. A system shown as an example has two inputs and one output where one of the inputs and the time derivative of the other are bounded. If the time optimum control is not unique, some other criterion must be added such that the maximum over or under swing of the transient, the area between the transient and the time axis, etc. are minimized. It is shown how to program an actual optimum control schedule. It is theoretically proved that all the transients based on the above control schedule are optimum and the theory is verified on an electronic analog computer.

BIBLIOGRAPHY

1. Oldenburger, R., "Optimum Nonlinear Control," Trans. ASME, Vol. 79, No. 3, pp. 456-527, (1957)

2. McDonald, D., "Nonlinear Techniques for Improving Servo Performance," Naional Electronics Conference, Vol. 6, pp. 400-421, (1950)

3. Hopkin, A. M., " A Phase-Plane Approach to the Compensation of Saturating Serco-Mechanisms," Trans. AIEE, Vol. 70, Part I, pp. 631-639, (1951)

4. Uttley, A. M. and Hammond, P. H., "The Stabilization of On-Off Controlled Servo-Mechanisms," Automatic & Manual Control, 1951 Cranfield Conference, pp. 285-307, (1952)

5. Flügge-Lotz, I., Discontinuous Automatic Control (book), Princeton University Press, Princeton, New Jersey, (1953)

6. Kazda, L.F., "Errors in Reley Servomechanisms," Trans. AIEE, Vol. 73, Part II, pp. 323-328, (1953)

7. Stout, T. M., "Effects of Friction in an Optimum Relay Servo-Mechanisms," Trans. AIEE, Vol. 73, Part II, pp. 329-336, (1953)

8. Silva, L. M., "Predictor Servomechanisms," Trans. IRE, Vol. CT-1, pp. 56-70, (1954)

9. Bogner, I. and Kazda, L.F., "An Investigation of the Switching Criteria for Higher Order Servomechanisms," Trans. AIEE, Vol. 73, Part II, pp. 118-127, (1954)

10. Chang, S. S. L., "Optimum Switching Criteria for Higher Order Contractor Servo with Interrupted Circuits," Trans. AIEE, Vol. 74, Part II, pp. 273-276, (1955)

11. Oldenburger, R., Nichlas, J.C. and Gamble, E.H., "Optimum Nonlinear Control of a Second Order Nonlinear System," National Aeronautics Space Administration Technical Note D-825, pp. 1-87, (1961)

12. Bushaw, D. W., "Optimum Discontinuous Forcing Terms," Contributions to the Theory of Nonlinear Oscillations, Princeton University Press, Vol. IV, pp. 29-52, (1958)

13. Bellman, R., Glicksberg, I. and Gross, O., "On the 'Bang-Bang' Control Probelm," Quarterly of Applied Mathematics, Vol. 14, pp. 11-18, (1956)

14. Pontryagin, L. S., "Some Mathematical Problems Arising in Connection with the Theory of Optimal Control Systems. Basic Problems of Automatic Regulation and Control," Proceedings of the Session of the AN SSSR on the Scientific Problems of the Automation of Production, (in Russian), Izd. AN SSSR, (1957)

15. Rozonoer, L. I., "L. S. Pontryagin's Maximum Principle in the Theory of Optimum Systems," Automation and Remote Control, Parts. I-III, Vol. 20, No. 10, pp. 1320-1334, 1405-1421, 1517-1532, (1959)

16. LaSalle, J. P., "The Time Optimal Control Problem," Contributions to the Theory of Nonlinear Oscillations, Pinceton University Press, Vol. V., pp. 1-24, (1960)

17. Friedland, B., "The Structure of Optimum Control Systems," Trans. ASME, Journal of Basic Engineering, Vol. 84, Series D, No. 1, pp. 1-11, (1962)

18. Markus, L. and Lee, E. B., "On the existence of Optimal Controls," Trans. ASME, Journal of Basic Engineering, Vol. 84, Series D, No. 1, pp. 13-20, (1962)

19. Rang, E. R., "A Switching Criterion for Certain Time-Optimal Regulating Systems," Trans. ASME, Journal of Basic Engineering, Vol. 84, Series D, No. 1, pp. 30-32, (1962)

20. Ho, Yu-Chi, "A Successive Approximation Technique for Optimal Control Systems Subject to Input Saturation," Trans. ASME, Journal of Basic Engineering, Vol. 84, Series D, No. 1, pp. 33-37, (1962)

21. Boyadjieff, G., Eggleston, D., Jaques, M., Sutabutra, H. and Takahashi, Y., "Some Applications of the Maximum Principle to Second Order Systems, Subject to Input Saturation, Minimizing Error and Effort," ASME Paper No. 62-WA-135, (1962)

APPENDIX

In deriving the first control function we assumed Relations (39) or (40). The reason for this is explained below. We consider only the case for $Y > 0$ since the treatment is identical for $Y < 0$.

Suppose that

$$Y > 0, \quad X_1 \geq -K \tag{48}$$

at $T = 0$. This condition yields the time T_{R2} negative where the time T_{R2} is given by Relation (30). The following cases may occur. A Φ trajectory obtained by setting $\Phi'' = 1$ touches or intersects the trajectory Δ_1 for $T < 0$ as shown in Figure 12 where the first case is represented by the trajectory Γ_2' and the second case by Γ_3'. The points at which the trajectory Γ_2' or Γ_3' has slope $-K$ is respectively on or below the trajectory Δ_1. Hence, if we substitute the values $X_1(0)$ and $Y(0)$ for the first case in Relation (38) we obtain

$$\Sigma_1(0) = 0$$

Similarly for the second case it follows that

$$\Sigma_1(0) < 0$$

However, the trajectories Γ_2' and Γ_3' lie above the trajectory Γ_2 for $T > 0$. Since our object in deriving the first control function is to know whether a Φ trajectory for $\Phi'' = 1$ is above, on or below the trajectory Γ_2 for $T = 0$, the above two cases are undesired results. This conflict may be overcome by the following procedure. We design the control schedule such that if Relations (48) are satisfied the optimal

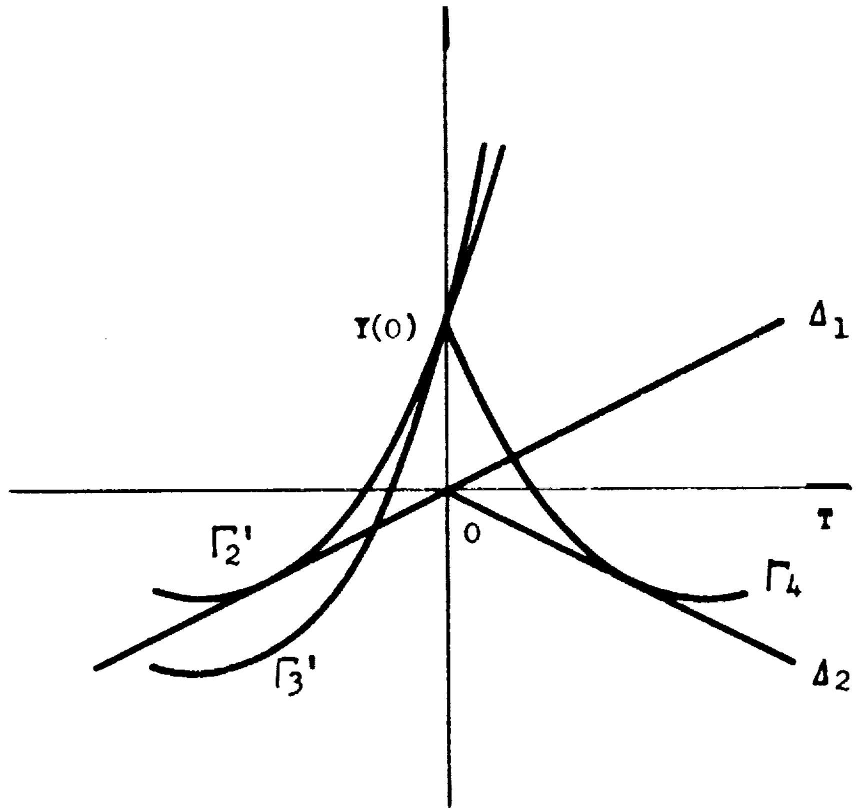


FIGURE 12 UNDESIRABLE CASES IN DERIVING FIRST CONTROL FUNCTION

selection of X_1' and X_2 is the same as for the case in which $\Sigma_1(0) > 0$ and Relations (39) simultaneously held. Then Transient Type B is obtained which is optimum. Similarly, we conclude that if

$$Y < 0, \quad X_1 \leq K \tag{49}$$

we let the optimal selection of X_1' and X_2 be the same as for the case in which $\Sigma_1 > 0$ and Relations (40) held.

VITA

Name:	Kiyohisa Okamura (Japanese Citizen)
Birth:	February 8, 1935, Chuseinan-do, Korea
Education:	B.S.M.E. Kyushu Institute of Technology, Japan, 1957
	M.S.M.E. University of Tokyo, Japan, 1959
	Purdue University, 1960 to 1963 - work toward Ph.D.
Experience:	Research Engineer, 1959 - 1960, Japan Atomic Energy Research Institute
	Research Associate, 1962 summer, Argonne National Laboratory